矩阵论选讲

管枫

七月在线

June, 2016

主要内容

- 线性空间与基
  - 举例说明线性空间的基本概念
- 线性映射与矩阵
  - 什么是矩阵?
  - 矩阵作为线性映射的代数表达方式
- 矩阵标准型
  - 使用矩阵有什么好处?
  - 矩阵的标准型理论可以帮助对线性映射进行分类
- 矩阵技巧及思想应用实例
  - 如何处理矩阵相关的问题?
  - 利用标准型理论，可以方便理解和解决很大一部分矩阵相关的问题

记号

- 本节课常用数学记号

V, W	向量空间
v, w	向量
$\mathbb{R}^n, \mathbb{R}^m$	实坐标空间
α, β	V 和 W 的基
$T: V \to W$	向量空间 V 到 W 的线性映射
$A_{\alpha,\beta}(T)$	线性映射 T 在 α 和 β 这两组基下的矩阵
$G(v_1, v_2)$	内积空间 V 上的内积
H_α	G 在基 α 下的矩阵形式

线性空间与基

实系数线性空间是一个由向量组成的集合, 向量之间可以做加减法, 向量与实数之间可以做乘法, 而且这些加, 减, 乘运算要求满足常见的交换律和结合律. 我们也可以类似地定义其他系数的线性空间.

Example (线性空间)

有原点的平面。

- 如果平面有一个原点 O, 那么平面上任何一个点 P, 都对应着一个向量 $\overrightarrow{OP}$。
- 这些向量以及他们的运算结构放在一起，就组成一个向量空间。

- 原点 O 在空间中引入了线性结构。(向量之间的加法，以及向量与实数的乘法)

线性空间与基

基是线性空间里的一组向量，使得任何一个向量都可以唯一的表示成这组基的线性组合.

Example (坐标空间)

有原点的平面，加上一组基 $\{\overrightarrow{X}, \overrightarrow{Y}\}$。

- 任何一个向量 $\overrightarrow{OP}$, 都可以唯一表达成 $\overrightarrow{OP} = a\overrightarrow{X} + b\overrightarrow{Y}$ 的形式。
- (a, b) 就是 P 点的坐标。

- 基给出了定量描述线性结构的方法——坐标系。

Example (坐标系的选择)

考虑纽约中城区的地图。街道用数字编号，但不是正南正北。在这个地图上以中心为原点，如何选取基?

- 基的选择取决于要解决的问题。
- 没有十全十美的基，只有适合解决问题的基。

小结 (线性空间与基)

- 线性空间是一种结构 *(加法及乘法运算结构)*
- 基使得我们可以用坐标描述线性结构
- 基的选择取决于要研究的问题

线性映射与矩阵

Definition (线性映射)

V 和 W 是两个实线性空间，$T: V \to W$ 如果满足如下条件就是一个线性映射。

$$
\begin{array}{lll}
(i) & T(v_1 + v_2) = T(v_1) + T(v_2), & \forall v_1, v_2 \in V \\
(ii) & T(\lambda v) = \lambda T(v), & \forall \lambda \in \mathbb{R}, v \in V
\end{array}
$$

- 线性映射的本质就是保持线性结构的映射
- 到自身的线性映射 $T: V \to V$ 叫做线性变换

线性映射与矩阵

线性变换的矩阵描述

V, W 分别为 n, m 维的线性空间，
$\alpha = \{\alpha_1, \dots, \alpha_n\}, \beta = \{\beta_1, \dots, \beta_m\}$ 分别为 V, W 的一组基。
$T: V \to W$ 是一个线性映射。于是 T, α, β 唯一决定一个矩阵 $A_{\alpha,\beta}(T) = [A_{ij}]_{m\times n}$，使得

$$T(\alpha_j) = \sum_{i=1}^{m} A_{ij} * \beta_i, \forall j \in 1, \dots, n \tag{1}$$

(1) 等价于

$$T(\alpha\ , \dots, \alpha_n) = (\beta_1, \dots, \beta_m) \cdot A_{\alpha,\beta}(T) \tag{2}$$

简记为

$$T(\alpha) = \beta \cdot A_{\alpha,\beta}(T) \tag{3}$$

如果我们选取 V, W 的另外一组基，$\tilde{\alpha} = \tilde{\alpha} \cdot P$, $\tilde{\beta} = \tilde{\beta} \cdot Q$. 那么存在矩阵 $A_{\tilde{\alpha},\tilde{\beta}}(T)$ 使得,

$$T(\tilde{\alpha}) = \tilde{\beta} \cdot A_{\tilde{\alpha},\tilde{\beta}}(T)$$

两边分别代入 $\tilde{\alpha}$ 与 $\tilde{\beta}$ 得到,

$$T(\alpha) \cdot P = T(\alpha \cdot P) = \beta \cdot Q \cdot A_{\tilde{\alpha},\tilde{\beta}}(T)$$

与(3)比较我们得到矩阵变换公式:

$$Q \cdot A_{\tilde{\alpha},\tilde{\beta}}(T) \cdot P^{-1} = A_{\alpha,\beta}(T) \tag{4}$$

小结 (线性映射与矩阵)

- 矩阵是线性映射在特定基下的一种定量描述

$$T(\alpha) = \beta \cdot A_{\alpha,\beta}(T)$$

- 基变换下的矩阵变换公式的推导方法

$$Q \cdot A_{\tilde{\alpha},\tilde{\beta}}(T) \cdot P^{-1} = A_{\alpha,\beta}(T)$$

线性映射与矩阵

例题 (几何变换：拉伸，反转与旋转)

实数平面 $\mathbb{R}^2$ 到自身的映射：

- 拉伸 $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
- 反转 $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
- 旋转 $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos(\pi/6) & \sin(\pi/6) \\ -\sin(\pi/6) & \cos(\pi/6) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Example (算法问题：如何计算斐波那契数列)

斐波那契数列：$a_1 = 1, a_2 = 1$，并且满足 $a_{n+2} = a_n + a_{n+1}$.

Solution (1. 简单递归)

我们可以使用定义进行简单递归

```
def arrayComputer(n):
    if n <= 2:
        return 1
    else:
        prevA = arrayComputer(n-1)
        prevprevA = arrayComputer(n-2)
        return prevA + prevprevA
```

Solution (2. 建立线性模型)

把 $[a_n, a_{n+1}]^T$ 看成一个向量，于是递归公式变成一个线性模型：

$$\begin{bmatrix} a_{n+1} \\ a_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}$$

于是：

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^{n-1} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

线性模型算法的 Python code:

```python
def linearComputer(n):
    if n <= 2:
        return 1
    current2A = [1, 1]
    prev2A = copy.deepcopy(current2A)
    for ind in range(n-2):
        current2A[0] = 0*prev2A[0] + 1*prev2A[1]
        current2A[1] = 1*prev2A[0] + 1*prev2A[1]
        prev2A = copy.deepcopy(current2A)
    return prev2A[1]
```

思考题

如何计算类似的数列 $a_1 = 1, a_2 = 1, a_3 = 1$, 并且 $a_{n+3} = a_n + 2a_{n+1} + 3a_{n+2}$?

矩阵的标准型: 方阵的相似变换

如果 $T: V \to V$ 是一个线性变换, 那么对于 V 的两组基 α 与 $\tilde{\alpha} = \alpha \cdot P$, 线性变换 T 的矩阵分别为

$$A_\alpha(T) \text{ and } A_{\tilde{\alpha}}(T) = P^{-1} \cdot A_\alpha(T) \cdot P$$

方阵的相似变换

- 如果两个方阵 A 和 $\tilde{A}$ 满足, $\tilde{A} = P^{-1}AP$. 那么这两个方阵就互为相似矩阵
- 相似矩阵的几何意义是同一个线性变换在不同的基下的表达形式
- 当研究对象是线性变换的时候, 我们只关心矩阵在相似变换下不变的几何性质。

矩阵的标准型

相似变换下不变的性质

本节列举一些相似不变量，稍后利用约当标准型来介绍他们的几何意义。

- 行列式 (det)

$$\begin{aligned}\det(P^{-1}AP) &= \det(P^{-1})\det(A)\det(P)\\ &= \det(P^{-1})\det(P)\det(A)\\ &= \det(A)\end{aligned}$$

- 迹 (trace),$\mathrm{tr}(AB) = \mathrm{tr}(BA)$

$$\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(APP^{-1}) = \mathrm{tr}(A\cdot I) = \mathrm{tr}(A)$$

- 秩 (rank)

矩阵的标准型: 相似不变量

相似变换下不变的性质

- 特征值：特征方程 $\det(A-\lambda I)=0$ 的根。
  如果 $\det(A-\lambda I)=0$，那么 $\det(P^{-1}(A-\lambda I)P)=0$, 于是 $\det(P^{-1}AP-\lambda I)=0$

- 特征值是最重要的相似不变量，利用这个相似不变量可以方便的得出上面所有的不变量。

矩阵的标准型：相似不变量

Theorem (矩阵的相似标准型)

任何一个实系数方阵 A, 都存在一个可逆实系数方阵 P, 使得 $P^{-1}AP$ 是一个分块对角矩阵 $diag(J_1, ..., J_k)$. 且每一个对角块 J_k 是如下四种情况之一：

i 1×1 伸缩变换矩阵块 $[\lambda]$

ii 2×2 伸缩旋转变换矩阵块 $R_{(\mu,\theta)} = \mu \cdot \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$

矩阵的标准型：相似标准型

iii $m \times m$ 循环伸缩矩阵块 $\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}$

iv $m \times m$ 循环伸缩旋转矩阵块

$$\begin{bmatrix} R_{(mu,heta)} & I_{2\times 2} & 0 & \cdots & 0 \\ 0 & R_{(mu,heta)} & I_{2\times 2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & R_{(mu,heta)} \end{bmatrix}$$

如果对角块中只有 [i] 和 [ii] 两种，那么这个矩阵称作可复对角化矩阵.

Proof.

I have discovered a truly marvellous proof of this, which this margin is too narrow to contain.– Fermat □

Theorem (几乎所有方阵都可复对角化)

对于任何一个实系数方阵 A, 都存在一个可复对角化的矩阵序列 $\{A_i\}_{i=1}^{\infty}$, 使得

$$\lim_{i\to\infty} A_i = A$$

Proof.

证明思路:

1. 如果一个矩阵的特征多项式没有重跟，那么这个矩阵一定可以复对角化.
2. 几乎所有的矩阵特征多项式都没有重跟.

$\square$

例题 (行列式 (det)：线性映射的体积膨胀系数)

如果 A 是线性变换 $T: V \to V$ 的矩阵， C 是 V 里边的立方体，那么：

$$Volume(T(C)) = \det(A) \cdot Volume(C)$$

Proof

因为行列式是相似不变量，不失一般性，可以假定 A 就是约当标准型.
证明分两步：先对可以复对角化的矩阵 A 进行证明，然后证明一般情况.
第一步：如果 A 是复对角化的矩阵，那么 $A = \mathrm{diag}(J_1, ..., J_k)$, 其中 J_i 要么是一维的伸缩变换矩阵块，要么是二维的旋转变换矩阵块. 所以可以进一步把 A 写成

$$A = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_p, R_{(\mu_1,\theta_1)}, ..., R_{(\mu_q,\theta_q)})$$

于是 $\det(A) = \lambda_1 \cdot \lambda_2 \cdots \lambda_p \cdot \mu_1 \cdots \mu_q = \prod_{i=1}^{p} \lambda_i \cdot \prod_{j=1}^{q} \mu_j^2.$

Continue Proof

另一方面 $T(C)$ 是由 C 在前 p 个维度上进行拉伸，而在后面的维度上进行二维拉伸及旋转得到的. 所以

$$\mathrm{Volume}(T(C)) = \prod_{i=1}^{p} \lambda_i \cdot \prod_{j=1}^{q} \mu_j^2 \cdot \mathrm{Volume}(C)$$

所以

$$\mathrm{Volume}(T(C)) = \det(A) \cdot \mathrm{Volume}(C)$$

Continue Proof

第二步：对一般的 A，存在一个可复对角化的矩阵序列 $\{A_i\}_{i=1}^{\infty}$，及其所对应的线性变换序列 $\{T_i\}_{i=1}^{\infty}$, 使得：

$$\lim_{i\to\infty} A_i = A \text{ 而且 } \lim_{i\to\infty} T_i = T$$

根据第一步的结论, 对于任何 i 我们有,

$$\mathrm{Volume}(T_i(C)) = \det(A_i)\cdot \mathrm{Volume}(C).$$

因为体积与行列式都是连续函数，我们得到

$$\begin{aligned}
\mathrm{Volume}(T(C)) &= \lim_{i\to\infty} \mathrm{Volume}(T_i(C)) \\
&= \lim_{i\to\infty} \det(A_i)\cdot \mathrm{Volume}(C) \\
&= \det(A)\cdot \mathrm{Volume}(C)
\end{aligned}$$

于是证毕.

例题 (迹 (tr)：*exp*(tr) 是线性映射 *exp*(*A*) 的体积膨胀系数)

$$\exp(tr(A)) = \det(\exp(A))$$

Proof

首先对可复对角化的矩阵进行证明. 为了简化步骤我们不加证明的使用

Proposition

如果 A 是可复对角化矩阵，则存在一个复系数可逆方阵 Q, 使得 $\tilde{A} = Q^{-1}AQ$ 是复系数对角矩阵.

$$\tilde{A} = diag(\lambda_1, ..., \lambda_n)$$

Continue Proof

于是 $\exp(\tilde{A}) = \mathrm{diag}(\exp(\lambda_1), \ldots, \exp(\lambda_n))$

$$\begin{aligned}
\exp(\mathrm{tr}(A)) &= \exp(\mathrm{tr}(\tilde{A})) \\
&= \exp(\sum_{i=1}^{n} \lambda_i) \\
&= \prod_{i=1}^{n} \exp(\lambda_i) \\
&= \det(\tilde{A}) \\
&= \det(A)
\end{aligned}$$

对于一般的 A, 可以模仿上一个证明中的第二步，留作思考题。

Continue Proof

于是 $\exp(\tilde{A}) = \mathrm{diag}(\exp(\lambda_1), \ldots, \exp(\lambda_n))$

$$
\begin{aligned}
\exp(\mathrm{tr}(A)) &= \exp(\mathrm{tr}(\tilde{A})) \\
&= \exp(\sum_{i=1}^{n} \lambda_i) \\
&= \prod_{i=1}^{n} \exp(\lambda_i) \\
&= \det(\tilde{A}) \\
&= \det(A)
\end{aligned}
$$

对于一般的 A, 可以模仿上一个证明中的第二步，留作思考题。

矩阵的标准型：相似不变量的几何意义

例题 (秩 (rank)：像空间的维数)

如果 A 是线性变换 $T: V \to V$ 在基 α 下的矩阵，那么

$$rank(A) = \dim T(V)$$

Proof

注意矩阵的秩并不是一个连续函数，所以对于这个问题我们不能仅仅考虑可复对角化的矩阵。我们必须直接考虑一般情况。
假定 $A = \mathrm{diag}(J_1, ..., J_K)$，其中每一个对角块都是约当标准型四中对角块里面的一种。于是我们得到空间 V 以及变换 T 的直和分解

$$V = V_1 \oplus \cdots \oplus V_k$$
$$T_i = T|_{V_i} : V_i \to V_i, \quad \text{for all } i \text{ from } 1 \text{ to } K$$

Continue Proof

所以
$$\dim T(V) = \sum_{i=1}^{K} \dim T_i(V_i)$$
而且
$$\mathrm{rank}(A) = \sum_{i=1}^{K} \mathrm{rank}(J_i)$$
所以要证明
$$\mathrm{rank}(A) = \dim T(V)$$
只需要对所有的约当块 J_i, 证明
$$\mathrm{rank}(J_i) = \dim T_i(V_i)$$
对于每一个约当块的证明留作思考题。

矩阵的标准型

小结 (矩阵的标准型)

- 任何一个方阵总存在约当标准型（线性变换的相关问题总是存在一组好基）
- 如果问题具有相似不变性，可以假定矩阵是约当标准型从而简化问题
- 如果问题具有相似不变性以及连续性，可以假定矩阵是可复对角化的约当标准型，从而进一步简化问题

Example (如何求斐波那契数列的通项公式)

斐波那契数列：$a_1 = 1, a_2 = 1$，并且满足 $a_{n+2} = a_n + a_{n+1}$.

Solution (回顾)

把 $[a_n, a_{n+1}]^T$ 看成一个向量，于是递归公式变成一个线性模型：

$$\begin{bmatrix} a_{n+1} \\ a_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}$$

于是：

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^{n-1} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Solution (利用标准型简化计算)

假设 $A=\begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}$, 约当标准型 $\tilde{A}=P^{-1}AP$, 那么

$$A=P\tilde{A}P^{-1}$$

带入到线性模型得到

$$\begin{bmatrix}a_{n+1}\\ a_{n+2}\end{bmatrix}=P\tilde{A}P^{-1}\cdot\begin{bmatrix}a_n\\ a_{n+1}\end{bmatrix}$$

于是

$$P^{-1}\begin{bmatrix}a_{n+1}\\ a_{n+2}\end{bmatrix}=\tilde{A}P^{-1}\cdot\begin{bmatrix}a_n\\ a_{n+1}\end{bmatrix}$$

矩阵技巧及思想应用实例

Solution (利用标准型简化计算)

所以

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = P\tilde{A}^{n-1}P^{-1} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

计算并带入 $\tilde{A} = \begin{bmatrix} \frac{1-\sqrt{5}}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{bmatrix}$ 以及 $P^{-1} = \begin{bmatrix} \frac{-1+\sqrt{5}}{2} & -1 \\ \frac{1+\sqrt{5}}{2} & 1 \end{bmatrix}$ 即可得所求.

思考题

如何计算序列 $a_{n+3} = a_n + a_{n+1} + a_{n+2}, a_1 = 1, a_2 = 1, a_3 = 1$ 的通项公式.

谢谢大家!